AF
2700

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

Application of: Donald J. Hejna, Jr.

Serial No.: 09/398,612 Filed: September 16, 1999

Examiner: A. A. Armstrong Group Art Unit: 2654

Title: **Method and Apparatus to Determine and Use Audience Affinity and Aptitude**

**SUBMISSION OF APPELLANT'S APPEAL BRIEF**



Mail Stop Appeal Brief-Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450

Sir:

Transmitted herewith is an original and two copies of Appellant's Appeal Brief (with a check of $165.00) and an Extension Petition for the above-identified patent application (with a check of $475.00).

Respectfully submitted,

Date: December 24, 2003 By Michael B. Einschlag

Michael B. Einschlag
Reg. No. 29,301
(650) 949-2267
25680 Fernhill Drive
Los Altos Hills, California 94024

Attached:

an original and two (2) copies of an Appeal Brief (and check of $165.00)
Extension Petition (and check of $475.00)

**CERTIFICATE OF MAILING under 37 C.F.R. 1.8(a)**

I hereby certify that this correspondence is being deposited on December 24, 2003 with the United States Postal Service as first class mail, with sufficient postage, in an envelope addressed to Mail Stop Appeal Brief-Patents, Commissioner for Patents, P.O. 1450, Alexandria, Virginia 22313-1450.

| Michael B. Einschlag | 29,301 | December 24, 2003 |
|---|---|---|
| Michael B. Einschlag | Reg. No. | Date of Signature |

#16

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

**BOARD OF PATENT APPEALS AND INTERFERENCES**

| | |
|---|---|
| **Applicant:** | **Donald J. Hejna, Jr.** |
| **Title:** | **Method and Apparatus to Determine and Use Audience Affinity and Aptitude** |
| **Serial No.:** | **09/398,612** |
| **Filed:** | **September 16, 1999** |
| **Examiner:** | **A. A. Armstrong** |
| **Group Art Unit:** | **2654** |
| **Date:** | **December 24, 2003** |



**APPEAL BRIEF**

**CERTIFICATE OF MAILING under 37 C.F.R. 1.8(a)**

I hereby certify that this correspondence is being deposited on December 24, 2003 with the United States Postal Service as first class mail, with sufficient postage, in an envelope addressed to Mail Stop Appeal Brief-Patents, Commissioner for Patents, P.O. 1450, Alexandria, Virginia 22313-1450.

Michael B. Einschlag — 29,301 Reg. No. — December 24, 2003 Date of Signature

01/02/2004 NDAMTE1 00000094 09398612

01 FC:2402 165.00 OP

(1) **Real Party in Interest.**

The real party in interest is Enounce Incorporated, the assignee of all right, title, and interest in and to the patent application.

(2) **Related Appeals and Interferences.**

There are no other appeals or interferences known to Appellant, Appellant's legal representative, or assignee which will directly affect or be directly affected by or have a bearing on the Board's decision in the pending appeal.

(3) **Status of Claims.**

Claims 1-13 are all the pending claims in the present patent application. Claims 1-13 have been rejected twice. Claims 1-13 are appealed.

(4) **Status of Amendment.**

No amendment was filed subsequent to the last rejection.

(5) **Summary of the Invention.**

The technology covered by the various claims relates in general to media works. As defined in the specification at p. 15, line 4-20: "A Media Work ("MW") may comprise, without limitation, one or more of text, pictures, audio, for example, a speech, an audio-visual work, for example, a movie or instructional video tape. ... In addition, an MW includes a collective MW which comprises a number of MWs. In further addition, an MW includes a MW created by combing an MW (a Target MW) and a set of reference information which can be used to reference portions of the Target MW. For example, the reference information may comprise hyperlinks to segments of an MW.

In addition, various aspects of the technology relate to presentation of such media works at various presentation rates. As set forth in the specification at p. 20, lines 5-7: "A Presentation Rate ("PR") comprises a information that can be used to obtain a rate at which a Media Work ("MW") is presented to an Audience." In particular, a playback where time advances at a rate that is different from the normal rate has a "presentation rate" that differs from the normal rate. For a playback where, for example, actions appear more rapidly (i.e., they occur at earlier times) than they actually occurred, the presentation rate is said to be increased. For a

playback where, for example, actions appear more slowly (i.e., they occur at later times) than they actually occurred, the presentation rate is said to be decreased.

Lastly, various aspects of the technology relate to novel concepts, i.e., audience affinity and audience aptitude. As set forth in the specification at p. 11, lines 14-17: "Audience Affinity Information ("AAffI") comprises an indicium of affinity of an Audience (defined, for example, by Audience interest or entertainment value to an Audience) for content properties, concepts, and the like." Further, as set forth in the specification at p. 11, lines 23-25: "Audience Aptitude Information ("AAptI") comprises an indicium of aptitude (defined, for example, by Audience familiarity or Audience fluency) with respect to content properties, concepts and the like."

In particular, claims 1-2 relate to methods for inferring audience affinity or aptitude; claims 3-4 relate to methods of utilizing audience affinity or audience aptitude; and claim 7 relates to a method of testing audience aptitude. Claim 5 relates to a method of presenting a media work by reordering depending on detected content or properties; and claim 6 relates to a method of presenting a media work in an order and at a presentation rate that depends on detected content or properties. Claims 8-9 relate to methods of presenting a media work by accessing information identifying the media work, a time to retrieve it, and a presentation rate. Claim 10 relates to a method of presenting a media work at presentation rates that depend on detected content properties, and presenting at a substantially uniform rate of content presentation. Claim 11 relates to a method of presenting a media work at presentation rates that depend on detected content properties which are indicia of actions of objects. Claims 12-13 relate to a method of determining a duration of a media work where presentation rates of various portions have been changed.

**The following briefly summarizes the invention of the various claims in light of the above.**

**Claim 1** relates to a method for inferring audience affinity or aptitude with regard to content or properties of portions of a media work which comprises: (a) presenting the media work to an audience; (b) obtaining user input regarding presentation rates for the portions of the

media work; (c) correlating the content or properties of the portions with the presentation rates; and (d) associating audience affinity or aptitude with the presentation rates for the correlated content or properties. As set forth in the specification at p. 23, lines 23-24: "... the manner in which the PRs are altered by the Audience members serve as a proxy for Audience affinity and/or Audience aptitude." In particular, as set forth in the specification at p. 23, line 27 to p. 24, line 6: "Advantageously, analyzing Audience input in accordance with the present invention to determine Audience affinity and aptitude, enables one to anticipate Audience response to previously unperceived MWs comprised of information and information properties to which Audience affinity and aptitude has been determined. This enables one to prepare information for use in presenting the unperceived MWs that will track Audience affinity and aptitude by causing the unperceived MWs to slow down and/or speed up in accordance with the analyzed affinity and aptitude." Please refer to the specification at p. 26, line 21 to p. 27, line 22; p. 27, line 24 to p. 21, line 21; p. 24, line 10 to p. 26, line 20 in conjunction with FIG. 22; p. 38, line 13 to p. 39, line 9 in conjunction with FIG. 3; and p. 98, line 25 to p.103, line 24.

**Claim 2** (depends from claim 1) relates to a method for inferring audience affinity or aptitude with regard to content or properties of portions of a media work which comprises: (a) presenting the media work to an audience; (b) obtaining user input regarding presentation rates for the portions of the media work; (c) correlating the content or properties of the portions with the presentation rates; and (d) associating audience affinity or aptitude with the presentation rates for the correlated content or properties; (e) wherein the presentation rates include a rate which causes a portion to be skipped. Please refer to the specification at p. 26, line 21 to p. 27, line 22; p. 27, line 24 to p. 21, line 21; p. 24, line 10 to p. 26, line 20 in conjunction with FIG. 22; p. 38, line 13 to p. 39, line 9 in conjunction with FIG. 3; p. 98, line 25 to p.103, line 24; and p. 56, line 19 to p. 57, line 9; and p. 102, lines 17-201; and p. 117, lines 6-13.

**Claim 3** relates to a method of utilizing audience affinity or aptitude associated with content or properties to present a media work which comprises: (a) detecting the content or properties in a portion of the media work; (b) associating the audience affinity or aptitude associated with the detected content or properties with a presentation rate for the portion; and (c)

presenting the portion at the presentation rate. Please refer to the specification at p. 77, line 11 to p. 81, line 24 in conjunction with FIGs. 20-21 and 24.

**Claim 4** (depends from claim 3) relates to a method of utilizing audience affinity or aptitude associated with content or properties to present a media work which comprises: (a) detecting the content or properties in a portion of the media work; (b) associating the audience affinity or aptitude associated with the detected content or properties with a presentation rate for the portion; and (c) presenting the portion at the presentation rate; (d) wherein associating includes accepting user input to determine the presentation rate. Please refer to the specification at p. 77, line 11 to p. 81, line 24 in conjunction with FIGs. 20-21 and 24.

**Claim 5** relates to a method of presenting a media work which comprises: (a) detecting content or properties in portions of the media work; (b) associating a presentation order with the detected content or properties that is different from the order of detection; (c) reordering the portions according to the presentation order; and (d) presenting the media work in accordance with the presentation order. Please refer to the specification at p. 88, line 13 to p. 92, line 3 in conjunction with FIG. 25.

**Claim 6** relates to a method of presenting a media work which comprises: (a) detecting content or properties in portions of the media work; (b) associating a presentation order with the detected content or properties that is different from the order of detection; and (c) presenting the media work in accordance with the presentation order; (d) wherein the step of associating further comprises associating a presentation rate of the portion with the detected content or properties; and the step of presenting comprises presenting the media work in accordance with the presentation order and the presentation rates. Please refer to the specification at p. 88, line 13 to p. 92, line 3 in conjunction with FIG. 25.

**Claim 7** of testing aptitude of an audience for content or properties of portions of a media work which comprises: (a) presenting the media to the audience; (b) obtaining user input regarding presentation rates for the portions of the media work; and (c) correlating the presentation rates with the aptitude for the content or properties of the portions. Please refer to the specification at p. 26, line 21 to p. 27, line 22; p. 27, line 24 to p. 21, line 21; p. 24, line 10 to

p. 26, line 20 in conjunction with FIG. 22; p. 38, line 13 to p. 39, line 9 in conjunction with FIG. 3; and p. 98, line 25 to p.103, line 24.

**Claim 8** relates to a method of presenting a media work having a presentation rate which comprises:: (a) accessing information identifying the media work and a time to retrieve the media work; (b) retrieving the identified media work at the time; (c) accessing presentation rate information to obtain a new presentation rate for use in altering the media work; and (d) altering the media work to create an altered work having the new presentation rate. Please refer to the specification at p. 5, lines 3-8; and p. 108, line 4 to p. 109, line 21.

**Claim 9** (depends from claim 8) relates to a method of presenting a media work having a presentation rate which comprises:: (a) accessing information identifying the media work and a time to retrieve the media work; (b) retrieving the identified media work at the time; (c) accessing presentation rate information to obtain a new presentation rate for use in altering the media work; (d) altering the media work to create an altered work having the new presentation rate; concatenating at least two altered media works to form a concatenated media work; and presenting the concatenated media work. Please refer to the specification at p. 5, lines 3-8; and p. 108, line 4 to p. 109, line 21.

**Claim 10** relates to a method of presenting a media work which comprises: (a) detecting media work content properties in a portion of the media work; (b) associating a presentation rate of the portion with the detected media work content properties; and (c) presenting the portion at the presentation rate; wherein the presentation rates provide a substantially uniform rate of content presentation. Please refer to the specification at p. 96, line 20 to p. 98, line 17.

**Claim 11** relates to a method of presenting a media work which comprises: (a) detecting media work content properties in a portion of the media work; (b) associating a presentation rate of the portion with the detected media work content properties; (c) presenting the portion at the presentation rate; and (d) wherein the media work content properties comprise indicia of actions of objects. Please refer to the specification at p. 96, line 20 to p. 98, line 17.

**Claim 12** relates to a method of determining the duration of an altered media work having a presentation rate of one or more of its segments that differs from that of a media

work used to create the altered media work, which method comprises: (a) segmenting the media work into segments having a single presentation rate; (b) determining the length of the segments of the media work; (c) computing the duration of the segments of the media work after application of the presentation rate; and (d) summing the durations to determine the duration of the altered media work. Please refer to the specification at p. 107, lines 11-27 in conjunction with FIGs. 29-30.

**Claim 13** (depends from claim 12) relates to a method of determining the duration of an altered media work having a presentation rate of one or more of its segments that differs from that of a media work used to create the altered media work, which method comprises: (a) segmenting the media work into segments having a single presentation rate; (b) determining the length of the segments of the media work; (c) computing the duration of the segments of the media work after application of the presentation rate; (d) summing the durations to determine the duration of the altered media work; and (e) excising segments from the media work having a presentation rate that exceeds a predetermined threshold. Please refer to the specification at p. 107, lines 11-27 in conjunction with FIGs. 29-30; and p. 105, line 16 to p. 106, line 17.

**(6) Issues.**

1. Whether claims 1-4, 7 and 10-11 are patentable under 35 U.S.C. § 103(a) over Richard et al. (U.S. Patent No. 5,924,068) in view of Oikawa et al. (U.S. Patent No. 5,396,577).
2. Whether claims 5-6 are patentable under 35 U.S.C. § 103(a) over Richard et al. (U.S. Patent No. 5,924,068) in view of Oikawa et al. (U.S. Patent No. 5,396,577) and well known prior art.
3. Whether claims 8-9 are patentable under 35 U.S.C. § 103(a) over Richard et al. (U.S. Patent No. 5,924,068) in view of Yumura et al. (U.S. Patent No. 5,752,228).
4. Whether claims 12-13 are patentable under 35 U.S.C. § 103(a) over Yumura et al. (U.S. Patent No. 5,752,228) in view of Richard et al. (U.S. Patent No. 5,924,068).

**(7) Grouping of Claims.**

Claims 1 and 7 stand or fall together on each ground of rejection.

Claim 2 stands or falls on its own on each ground of rejection.

Claim 3 and 4 stand or fall together on each ground of rejection.

Claim 5 stands or falls on its own on each ground of rejection.

Claim 6 stands or falls on its own on each ground of rejection.

Claim 8 stands or falls on its own on each ground of rejection.

Claim 9 stands or falls on its own on each ground of rejection.

Claim 10 stands or falls on its own on each ground of rejection.

Claim 11 stands or falls on its own on each ground of rejection.

Claim 12 stands or falls on its own on each ground of rejection.

Claim 13 stands or falls on its own on each ground of rejection.

The reasons that the above-identified groups of claims stand or fall together are included in the appropriate portions of the "Argument" section of this brief.

**(8) Argument.**

**Issue 1: Whether claims 1-4, 7 and 10-11 are patentable under 35 U.S.C. § 103(a) over Richard et al. (U.S. Patent No. 5,924,068) in view of Oikawa et al. (U.S. Patent No. 5,396,577).**

**Reasons why claims (1 and 7) and 2, (3 and 4), 10 and 11 are separately patentable.**

Appellant respectfully submits that claim 2 is separately patentable because claim 2 further comprises presentation rates that "include a rate which causes a portion to be skipped." In addition, Applicant respectfully submits that claims 3 and 4 are separately patentable because claims 3 and 4 relate to a method of utilizing audience affinity or aptitude whereas claims 1-2, and 7 relate to methods of inferring and testing, respectively, audience affinity or aptitude; and claims 10-11 relate to methods of presenting a media work. In further addition, Applicant respectfully submits that claim 10 is separately patentable because claim 10 comprises presenting a media work wherein the presentation rates provide a substantially uniform rate of content presentation. Lastly, Applicant respectfully submits that claim 11 is separately patentable because claim 11 comprises presenting a media work by associating a presentation rate of a portion of a media work with content properties that comprise indicia of actions of objects.

**Richard et al.**

As set forth in the Abstract, Richard et al. discloses: "An electronic news receiving device receives text data for an electronic edition of a newspaper in the evening and audibly reads the newspaper to the user the next day. ... The received electronic edition of the newspaper is processed by a section filter to retain desired sections of the newspaper and to discard unwanted sections. The retained news articles are stored in memory. A text-to-speech converter produces an audible output corresponding to the spoken text of the news articles. A user can input one or more keywords to cause the device to selectively read articles containing the keywords. The text to speech converter of the device uses rules and a dictionary to provide syntactic and semantic prosody for morpheme reconstructions. The user may determine which articles are read and may vary the rate at which articles are read using manual controls or spoken commands."

As set forth at col. 2, lines 60-63, Richard et al. teaches that: "Receiver 130, located within the newsreader 100, receives the marked news articles (marked so the articles can be interpreted) transmitted by the transmitter 120. (Emphasis added)" Next, at col. 2, line 66 to col. 3, line 7, Richard et al. teaches that: "A section filter 140 retains sections of the news that interest the user and discards the remaining sections. The section filter 140 performs a search algorithm, which is set by the user, to look for articles of certain types. ... The section filter 140 includes a memory 141 that stores a list of sections selected by the user. The articles that pass through the section filter 140 are sent to a linked list generator 145 which stores the desired articles in a low capacity memory 150." Next, at col. 3, lines 36-39, Richard et al. teaches that: "User interface 160 is used to establish the search algorithm performed by the section filter 140. The user interface 160 is also used to control playback of stored news article during playback mode." Next, at col. 3, lines 41-47, Richard et al. teaches that: "A text-to-speech converter 170 and a speaker 180 are coupled to the user interface 160. The news articles stored in the low capacity memory 150 and retrieved through user interface 160 are provided to the text-to-speech converter 170. The text-to-speech converter 170 and speaker 180 provide an audible output signal which corresponds to the text of news articles." Next, at col. 6, lines 26-31, Richard et al. teaches that: "User interface 160 can be one of two types. The basic user interface 160 allows

only a limited number of operations and retrieves news articles based on the section headings. The advanced user interface 166 (shown in FIG. 19) provides keyword searching and other advanced news article retrieval operations." Next, at col. 7, lines 8-47, Richard et al. teaches that:

> The setup controls 164 and the playback controls 162 make up the basic user input interface 160. ... To set-up the newsreader 100, the user first presses the Set-up/playback button 524 to enter the set-up mode. ... Entering the set-up mode causes a list of available news sections to be retrieved from memory 141 (shown in FIG. 1). ... The available sections are transmitted to the newsreader when a subscriber first requests delivery of the electronic edition of the newspaper.
>
> The section filter 140 includes a memory 141 (shown in FIG. 1) which stores the available sections. The user can scroll through the sections ... . If the user wants the section filter 140 to retain the articles in a section, the Select button 523 is pressed while the section title is displayed on the LCD 430. ... .
>
> ... When the user has completed selecting the desired sections, the Set-up/playback button 524 is pressed. This exits set-up mode and stores the selected section information in memory 141 (shown in FIG. 1). The stored selected section information is used by section filter 140 to determine which sections of the received newspaper should be stored into the memory 150.

Next, at col. 8, lines 4-23, Richard et al. teaches that:

> Referring to FIG. 1, once the section list and the associated section bits have been stored in memory 141, the newsreader 100 is ready to receive the electronic newspaper. In the exemplary embodiment, the receiver 130 receives the electronic edition of the newspaper in the early morning hours before dawn. The receiver 130 contacts the transmitter 120 and initiates the transfer of the electronic edition of the newspaper to the newsreader 100. FIG. 9 shows hardware elements within the receiver 120 which are used for the automatic downloading of the electronic edition of the newspaper. As shown in FIG. 9, the receiver 130 includes a timer 1310. The timer 1310 determines whether it is time to call the electronic news preparer. At the appropriate time (some time in the early morning) the timer 1310 instructs the modem 1320 to call

> the transmitter. As described above, other communications devices may be used. For example, if the electronic edition of the newspaper is received through AM/FM SCA broadcast, the timer would enable a tuner/demodulator circuit (instead of modem 1320) to begin receiving the electronic edition of the newspaper. (Emphasis added)

Next, at col. 9, line 62 to col. 10, line 14, Richard et al. teaches that: "To playback the stored articles, the user operates the playback controls 162, shown in detail in FIG. 13. The user first presses the Sections button 511. The headlines of the articles in the current section are consecutively read to the user. In addition, the current section is displayed on LCD 430 (shown in FIG. 4). Once the headlines for the current section are read, the newsreader automatically begins reading the headlines for the next section. If the user wishes to switch sections, the Section button 511 is pressed. If the user wants to hear the entire article corresponding to a read headline, the Read button 512 is pressed. The newsreader 100 then reads the entire article. The newsreader 100 can also pause between each read headline to allow the user adequate time to determine if she wants to hear the article corresponding to the headline just read. The next headline of the next article in the section is read if the read button is not pressed during the pause. If the user does not desire to hear the entire article, the Skip button 513 is pressed, and the newsreader 100 resumes reading article headlines starting with the article immediately following the just-read article." Next, at col. 12, lines 59-64, Richard et al. teaches that: "The text-to-speech converter 170 (shown in FIG. 1) converts unrestricted text into a synthetic speech waveform. As shown in FIG. 18, there are three major components in the text-to-speech converter: a language analysis component 710, an acoustic processing component 720 and a synthesis component 730." Next, at col. 13, lines 3-16, Richard et al. teaches that: "Unrestricted text contains abbreviations, numbers, and special symbols. Thus, in pre-processing module 711 the input text is pre-processed in order to normalize the input text. For example, numerals can be converted from numeric form to word form. Common abbreviations may be expanded, i.e. Mr. to Mister. In addition to text normalization, different punctuation marks such as commas, periods, question marks, and colons are interpreted. For example, punctuation can be converted to a data element that represents a delay and an inflection change on the preceding word or

words. For example, in the sentence "The President went to Wyoming ?", the pitch of the word Wyoming would be raised by pitch calculation module 724 to indicate that this phrase was a question." Next, at col. 13, lines 21-33, Richard et al. teaches that: "Once the pre-processing is completed, the system extracts words from the input text character strings in dictionary search module 712. A dictionary is searched for the pronunciation and part-of-speech information of these words. To reduce memory requirements, the dictionary consists of syntactical units called morphemes, their pronunciations and parts of speech. A morpheme by itself may not actually be a word; however, when combined with certain prefixes or suffixes, it becomes a word. For example, the words "optimal", "optimize", "optimization", "optimist", and "optimism" are all derived from the morpheme "optim" in the lexicon and various suffixes." Next, at col. 13, lines 37-47, Richard et al. teaches that: "In the grammatical parse module 713, a parser uses grammatical rules to assign parts of speech to the input text and to determine phrase boundaries. This information is used in the duration calculation module 723 and the pitch calculation module 724 to produce more natural sounding prosody. In many cases, the grammatical parse module 713 can determine the intended part-of-speech of a homographic word, leading to the correct pronunciation by the synthesizer. For example, the word "contract" in the phrase "to contract" is interpreted by the parser as a verb, while in the phrase "a contract", it is interpreted as a noun." Next, at col. 13, lines 50-58, Richard et al. teaches that: "The input text is then sent to the morphophonemic/letter-to-sound module 714 where letter-to-sound rules provide pronunciations for those words not derived from the dictionary search module 712. In addition, the words that are derived from morphemes in the dictionary may need pronunciation and stress changes at morpheme boundaries. For example, in the work "treated", derived from "treat"+"ed", the pronunciation of the "ed" ending must be changed from /D/ to /ID/." Next, at col. 14, lines 9-39, Richard et al. teaches that: "The acoustic processing component 720 uses rule-based systems to (1) modify basic word pronunciation according to grammatical information and phonemic context, (2) assign durations for segments and intonation patterns for intonational phrases, and (3) convert phonemes and features to acoustic parameters. The word-level stress assignment module 721 is responsible for assigning stress to each word based on its part of speech and number of syllables. The word-level stress assignment module 721 calculates the amount of

stress to be assigned to each word, using the information supplied by the language processing component 710. This stress information is then used by the duration calculation module 723, pitch calculation module 724, and phonetic translation module 725." Next, at col. 14, lines 36-43, Richard et al. teaches that: "The duration calculation module 723 computes the length of each phoneme segment based on several observations about the segment and its environment. For instance, vowel phonemes are generally longer than consonant phonemes. Also, phonemes which precede the vowel in a stressed word are longer than the same phonemes in a non-stressed word." Next, at col. 14, lines 54-61, Richard et al. teaches that: "After the input text is processed by the phonological translation module 722, duration module 723, and pitch module 724, the phonetic module 725 produces a set of synthesis parameters for the text. ... The parameters are then sent to the synthesis component 730 to synthesize speech sounds." Next, at col. 15, lines 19-39, Richard et al. teaches that: "The advanced user interface 166 operates in a set-up mode which is similar to the set-up mode described above with respect to the basic user interface 160. ... During playback, the advanced user interface 166 allows simplified user entry of queries and the audio playback of selected news articles based on the keywords they contain." Next, at col. 15, lines 36 to 63, Richard et al. teaches that:

> This variation of the newsreader 100 allows retrieval of articles based on a keyword search. Of course, one of the keywords could be the section heading name so that all the articles in one section are retrieved as described above. The advanced user interface allows the user to retrieve all articles dealing with a particular topic regardless of the section headings. For example, a user could enter "corn" and receive articles from the Food section with the latest recipes and articles from the Business section discussing commodities futures. The normal sequence for selecting articles for playback is the following.
>
> 1. Type keywords to search for in the box. ...
> 2. Click on the Find button to start the retrieval.
> 3. When the results of the query are returned by the index engine (shown in FIG. 26), the headlines of each retrieved article are recited. At this time the Stop, Next, Previous, and Wait buttons are enabled. To listen to an article, press any key after hearing the headline.

4. Press Next, Previous, Wait/Continue, Stop, or Exit to control the playback. Next is enabled as long as there is at least one more article in the list of retrieved files. When pressed, either during headline or article recitation, the next headline is played back, and playback continues with headlines.

Next, at col. 16, lines 41-59, Richard et al. teaches that:

To provide for keyword searching, the newsreader 100 equipped with the advanced user interface 166 also includes an index engine 900 as shown in FIG. 22. Section filter 140 is the same as that shown in FIG. 11. The linked list generator 145 (shown in FIG. 11) is replaced with an index engine 900. The index engine 900 stores the articles in memory 150 so that keyword retrieval is possible. Each article is stored as a file. A keyword index, based on the words appearing in the articles, is established by the index engine 900. The advanced user interface 166 accesses the articles in memory 150 through the index engine 900.

FIG. 23 illustrates how the news articles, stored in the memory 150, may be indexed by the index exemplary engine 900. Index engine 900 maintains a hash table 950 that has as entries keywords derived from each news article. Each entry of the hash table points to a bucket 960. The bucket 960 includes a list of article entries, each entry contains a pointer to the article and position information within the article for the keyword.

Next, at col. 17, lines 37-49, Richard et al. teaches that: "The results from the index engine consist of a list of file names, each of which matches the keyword set in the query. The advanced user interface 166 (shown in FIG. 22) then reads each file in turn searching for its headline by locating the headline marker as shown in Table I. This headline string is in turn passed to the speech synthesizer utilizing the included application interface (API) where interprocess communications with the text-to-speech converter 170 take place. The process of reading of files, searching for the headline, passing the string to the synthesizer, and waiting to see if the user presses a key continues until the list of files is completed, at which time the newsreader returns to its initial state waiting for a query, or for a key to be pressed." Next, at col. 18 to col. 19, line 12, Richard et al. teaches that:

> If the newsreader 100 is equipped with the advanced user interface 166 described above, it may be used to control the various parameters of the text-to-speech converter 170. In the stand alone product shown in FIG. 4, these parameters are set by the manufacturer. Because the advanced user interface 166 discussed above allows a wider variety of user inputs, the following text-to-speech converter 170 parameters can be altered by the user.
>
> Voice Type
>
> The user may select between the male voice and the female voice by using a simple toggle command. Both voices in the default setting have a slightly low overall pitch which is designed to minimize the user's fatigue in listening to long text passages. This voice characteristic is desirable for reading and proofreading long text passages.
>
> Voice Pitch
>
> Once a voice type is selected, the user can make fine adjustments in pitch by raising or lowering the overall pitch for that voice.
>
> Speech Rate
>
> The text-to-speech converter 170 has an average default speech rate is 170 words per minute. The user can select an average speech rate from 120 to 240 words per minute.

Lastly, at col. 20, lines 34-40, Richard et al. teaches that: "The invention allows for delivery of an electronic edition of the newspaper to a subscriber in the evening and for audible playback of the newspaper the next day while the user is commuting to work or performing other tasks. By using either a basic or advanced user interface, the user can retain the news articles that are of interest and discard unwanted material."

**Oikawa et al.**

As set forth in the Abstract, Oikawa et al. discloses: "In a speech synthesizing apparatus, importance degree information indicative of a degree of importance with respect to each text portion of input original text data is added to this text portion. Then, the original text data with such importance degree information is input. When a rapid reading process, or a head searching process is carried out for the original text input, speech synthesis is carried out by controlling several stages which text portion should be skipped, or at which speed, the text

portions should be synthesized, in response to a speed instruction and importance degree information which are being input into the speech synthesizing apparatus." Next, at col. 2, lines 13-54, Oikawa et al. teaches that:

> The present invention ... has an object to provide such a speech synthesizing apparatus capable of performing a rapid reading process and a search process at a higher speed than that of the conventional speech synthesizing system, without increasing the overall system scale.
>
> To achieve the above-described object, the speech synthesizing apparatus 11 of the present invention, records input text data TX, which contains both input text data and information which describes the degree of importance with respect to each text portions.
>
> The speech synthesis process is carried out by skipping the text portions TX1, TX2, - - - , having a low degree of importance based upon the importance degree information previously recorded.
>
> Furthermore, the above-described speech synthesis apparatus 11 includes an input means 13 for designating synthesizing speed information 12G, which allows having a low degree of importance to be skipped during the speech synthesis process.
>
> In accordance with the present invention, since the importance degree information IP1, IP2, - - - , has been added to the respective text portions TX1, TX2 of the text data TX, the respective text portions TX1, TX2, - - - , of the relevant text data TX are categorized by levels indicative of the degrees of importance related to the relevant text portions TX1, TX2, - - - . This is required to facilitate the rapid reading process and the search process. As a consequence, one level of the multiple levels is designated in accordance with the speeds of the rapid reading process and of the search process, so that only such text portions TX1, TX2, - - - , having the same degree of importance may be disconnected and synthesized with each other while skipping nonsimilar text portions. Therefore, the rapid reading speed and the search speed of the present invention can be further increased, as compared with those of the conventional speech synthesizing system.

Next, at col. 3, lines 16-59, Oikawa et al. teaches that:

In the speech synthesizing apparatus 11 shown in FIG. 2, a text portion selecting unit 12 is provided at a prestage of the sentence analyzing unit 2, and a speed instruction generating unit 13 is externally employed. Then, as shown in FIG. 3A, a text portion corresponding to a skip level designated by a reading speed instruction is designated based upon degrees of importance for the text portions TX1, TX2, - - - , with employment of importance degree information IP1, IP2, - - - . The importance degree information has been inserted as information used to a head search, into head portions of the text portions TX1, TX2, - - - , of the input original text data TX. Accordingly, the process for designating the reading speed is executed.

It should be noted that the inserted importance degree information represent levels with respect to the degrees of importance about the subsequent text portions TX1, TX2, - - - , depending upon the contents thereof. ...

The text portion selecting unit 12 enters an input text-12A constructed of the original text data TX (see FIG. 3A) into a text analyzing block 12B. The text analyzing block 12B separates the original text data TX into the text portions TX1, TX2, - - - , and also the importance degree information IP1, IP2, - - - . The separated text portions 12C (i.e., symbols TX1, TX2, - - - , of FIG. 3A) are input into a reading segment selecting block 12D. On the other hand, the importance degree information 12E (namely, symbols IP1, IP2, - - - of FIG. 3A) is input into a reading segment determining block 12F, so that a determining process of a reading segment is executed at a speed defined by the speed instruction given from the speed instruction generating unit 13.

As a consequence, a reading instruction 12G produced by the reading segment determining block 12F contains instructions as shown in Table 1. That is, the text portions are eventually selected in the disconnected form, and simultaneously the text portions which are not read are skipped by selecting only the reading sections designated among the text portions TX1, TX2, - - - .

Next, at col. 4, lines 11 to col. 5, line 4, Oikawa et al. teaches that:

In this preferred embodiment, the skip levels "0," "1," and "2" defined in Table 1 are preset as follows: At the skip level "0", as shown in FIG. 3B, all of the text portions having the values of the importance degree information of "0," "1" and "2" are read. At the skip level "1," as indicated in FIG. 3C, the text portions having

the values of the importance degree information greater than "0" (namely, exclude the value of 0 are read. Further, at the skip level 2, as represented in FIG. 3D, the text portions with the values of the importance degree information larger than "1" (namely, exclude the values of "0" and "1") are read. Finally, as indicated in FIG. 3E, when the skip level becomes "3," the text portions with the values of the importance degree information greater than "2" (namely, exclude the values of "0," "1," "2") are read.

There are prepared three different sorts of the reading speeds, i e. "normal speed," "rapid speed 1," and "rapid speed 2."

...

In the speech synthesizing apparatus 11 with the above-described arrangement, as illustrated in FIG. 3A, the original text data TX used in the input text block 12A previously contains the importance degree information IP1, IP2, - - - , indicative of the importance degree (for example, the importance degree as the keyword) with respect to a series of text portions TX1, TX2, - - - . Then, the importance degree information IP1, IP2, - - - , 12E is separated from the text portion 12C by executing the process of the text analysis block 12B.

As a result, a series of importance degree information IP1, IP2, - - - which has been extracted, or separated from the original text data, is processed by the extracting process in the reading segment determining block 12F based on the skip levels indicated by the speed instructions issued from the speed instruction generating unit 13. Thus, the reading instruction 12G to designate the text portion to be read is produced by utilizing the extracted result.

Accordingly, the following selecting process is executed by the reading segment selecting block 12D. That is, as represented in FIGS. 3A to 3E, in accordance with the contents of the speed instruction issued from the speed instruction generating unit 13, when the skip level "0" is designated, all of the text portions are read. Similarly, when the skip level 1 is designated, the text portions with the importance degree information greater than 1 are read; when the skip level 2 is designated, the text portions with the importance degree information greater than 2 are read; and when the skip level 3 is designated, the text portions with the importance degree information greater than 3 are read. As a consequence, a series of text portions which have been selected in accordance with the skip levels are supplied to the text input block 2A of the sentence analyzing unit 2.

Next, at col. 5, lines 5-37, Oikawa et al. teaches that:

The sentence analyzing unit 2 analyzes the selected text portions to detect the words, boundaries of phrases, and basic accents in a similar manner to that of FIG. 1, on the basis of the dictionary (FIG. 2D).

The detection results of the words, boundaries of phrases, and basic accents are processed in accordance with a predetermined phoneme rule in the speech synthesizing rule unit 3, and then a synthesized parameter indicating when the text to be read under no intonation is produced. At this time, lengths of time for the respective phoneme are controlled in accordance with the speeds of the speed instructions so as to be coincident with the "normal reading" the "rapid reading 1" and the "rapid reading 2".

Furthermore, the detection results of the words, the boundaries of phrases, and the basic accents are processed in the speech synthesizing rule unit 3 in accordance with a predetermined phoneme rule in a similar manner to those of FIG. 1, so that a basic pitch pattern indicative of the intonation of the overall text input is produced in accordance with the speeds of the speed instructions.

...

With the above-described arrangement, according to the speech synthesizing apparatus 11, synthesized speech can be outputted when the input text is rapidly read, or read under skip condition in conformity to the speed instruction designated by the importance degree information contained in the input text.

Next, at col. 5, lines 38-55, Oikawa et al. teaches that:

Therefore, according to the speech synthesizing apparatus of the above-described arrangement, there are specific advantages when text to which the importance degree information has been added is speech-synthesized during rapid reading. For instance, in text which has been recorded on a medium, the structure of the original text data to be inputted (namely, a series of symbol containing information about words, boundaries of phrase, reading and basic accents), obtained by and analyzed in a sentence analyzing apparatus has been previously known. In this case, since several stages of the search levels can be set first, the capability to perform a search operation is increased. Secondly, since the head searching information, i.e., the importance degree information codes are contained in the input text, there is another advantage

that no care is taken to consider the head searching operation at the system side.

Next, at col. 6, lines 1-12, Oikawa et al. teaches that: "As previously described in detail, in accordance with the present invention, such a speech synthesizing apparatus for synthesizing speech from the input text can be readily realized, which processes and enters text after the importance degree information, indicative of the importance degree for the text portions, has been added thereto. When either the rapid reading process, or the head searching process is carried out, the speech can be synthesized while controls at several stages determine which text portions are skipped, or at which speed, the text portions are synthesized based on the speed instruction and the importance degree information."

An interview was held at the USPTO on October 8, 2003 among Examiner Armstrong, Inventor Hejna and Attorney Einschlag. Attorney Einschlag argued that Importance Level is independent of speed, and as support for that argument, pointed how Oikawa et al. teaches that each Importance Level can be played at a number of speeds (for example, see Table 1 at cols. 3-4). In an "Interview Summary" (PTOL-413) filled out by Examiner Armstrong, Examiner Armstrong stated "Oikawa the importance level does not effect the presentation rate."

**Combination of Richard et al. and Oikawa et al.**

Appellant respectfully submits that there is no reason, suggestion, or motivation in Richard et al. or Oikawa et al. or anywhere else that would have led one of ordinary skill in the art to combine Richard et al. and Oikawa et al to provide the methods of claims 1-4, 7, and 10-11. In particular, as discussed above, Richard et al. relates to an apparatus that receives an electronic version of a newspaper, stores pre-selected types of articles, "speaks" the headlines of stored articles, and "speaks the articles in response to user input. This is contrasted with Oikawa et al. which relates to a speech synthesis apparatus that "speaks" input text. To utilize the apparatus disclosed in Oikawa et al., importance degree information is added to the original text. Then, when a rapid reading process to carried out, a speed instruction and importance degree information is input, and speech synthesis is carried out by controlling which portions of the text should be skipped or at which speed the text portions should be synthesized. Applicant

respectfully submits that there is no reason to combine these references because they are so different. In particular, Richard et al. teaches an elaborate methodology for a user to determine which selections of an electronic newspaper are to be skipped by listening to headlines for various sections of the electronic newspaper, whereas Oikawa et al. requires pre-entering importance information. Further, the speed at which the newspaper text of Richard et al. is "spoken" is determined by using an advanced user interface whereas this is determined in Oikawa et al. by input speed and skipping level information that cannot be changed in real time.

In addition, Applicant respectfully submits that the Examiner has not provided the type of evidence of a teaching, motivation, or suggestion to combine these references that is required in these circumstances (namely, a rejection based on obviousness), see In re Sang Su Lee, (Fed. Cir. 00-1158, Decided Jan. 18, 2002) which states at p. 7, "When patentability turns on the question of obviousness, the search for and analysis of the prior art includes evidence relevant to the finding of whether there is a teaching, motivation, or suggestion to select and combine the references relied on as evidence of obviousness." and "'The factual inquiry whether to combine references must be thorough and searching.' ... It must be based on objective evidence of record."

Specifically, the Examiner's evidence regarding a teaching, motivation, or suggestion for combining Richard et al. and Oikawa et al. to provide the methods of claims 1-4, 7, and 10-11 is: "Therefore, it would have been obvious to one of ordinary skill at the time of invention to modify the system of Richard and implement associating playback rates based on specific categories as taught by Oikawa et al., for the purpose of ensuring that a user's preference for playback rates for a specific category of newspaper article is always maintained. (Emphasis added)" Applicant respectfully submits that this is a conclusory statement that does not fulfill the requirement (as set forth above) of the Federal Circuit. Further, because there is no teaching, motivation, or suggestion to combine Richard et al. and Oikawa et al. as asserted by the Examiner, Applicant respectfully submits that the Examiner's reasoning is based on improper hindsight. Still further, as will be set forth in detail below, even if one of ordinary skill in the art were to combine the teachings of Richard et al. and Oikawa et al., that one would not arrive at the methods of claims 1-4, 7 and 10-11.

**Regarding claims 1 and 7.** Applicant respectfully submits that a combination of the teachings of Richard et al. and Oikawa et al. does not render claims 1 or 7 obvious because there is no teaching, motivation, or suggestion anywhere to combine them to provide the inventions of claims 1 or 7. In addition, even if one combined Richard et al. and Oikawa et al., one would not arrive at the invention of claims 1 or 7.

In particular, claim 1 relates to a method for inferring audience affinity or aptitude and claim 7 relates to a method of testing aptitude of an audience for content or properties of portions of a media work. In accordance with claim 1 audience affinity or aptitude with regard to content or properties of portions of a media work is associated with presentation rates for correlated content or properties in the portions of the media work, and in accordance with claim 7, presentation rates for portions of the media work are correlated with the aptitude for the content or properties of the portions. There is nothing in Richard et al. or Oikawa et al. alone or in the combination of Richard et al. with Oikawa et al. that: (a) relates to audience affinity or aptitude; (b) relates to associating audience affinity or aptitude with presentation rates for correlated content or properties; or (c) relates to inferring or testing audience affinity or aptitude. As set forth in the specification at p. 11, lines 14-17: "Audience Affinity Information ("AAffI") comprises an indicium of affinity of an Audience (defined, for example, by Audience interest or entertainment value to an Audience) for content properties, concepts, and the like." Further, as set forth in the specification at p. 11, lines 23-25: "Audience Aptitude Information ("AAptI") comprises an indicium of aptitude (defined, for example, by Audience familiarity or Audience fluency) with respect to content properties, concepts and the like."

Richard et al. teaches "speaking" portions of a newspaper where headlines of articles are read to a user, and the user can elect to have the entire article read. Richard et al. also teaches that the user can cause a speech rate of the article to be changed. As the Examiner can readily appreciate from this, Richard et al. does not teach, hint or suggest, in any manner whatsoever, correlating content or properties of portions of the media work with speech rates input by the user. Further, Richard et al. does not teach, hint or suggest, in any manner whatsoever, inferring audience affinity or aptitude for portions of the newspaper. Still further, Richard et al. does not teach, hint or suggest, in any manner whatsoever, associating audience

affinity or aptitude for portions of the newspaper with the speech rates input by the user with content or properties. In other words, Richard et al. does not teach, hint or suggest, in any manner whatsoever, capturing information about the content or properties the user is listening to and a corresponding speed it is being listened to associate or correlate an audience affinity or aptitude.

Oikawa et al. teaches inputting text data and degree of importance data into a speech synthesis apparatus, and having the apparatus use the degree of importance data to determine whether to skip portions of the text. Oikawa et al. does not teach, hint or suggest, in any manner whatsoever, obtaining user input regarding presentation rates for a portion of the input, or correlating content or properties of the portion with the user input presentation rates. Still further, Oikawa et al. does not teach, hint or suggest, in any manner whatsoever, associating audience affinity or aptitude with the presentation rates correlated content or properties.

Applicant respectfully submits that Oikawa et al. teaches away from analyzing content of text data to perform rapid reading, see Oikawa at col. 1, line 62 to col. 2, line 10. In particular, Oikawa et al. teaches inputting data into a speech synthesis apparatus, which data contains both an input text portion and information which describes a degree of importance of the text portion. Oikawa et al. then teaches using the degree of importance to determine whether to skip portions of the text associated with degrees of importance below a selected amount. However, as disclosed at col. 3, line 37 to col. 4, line 31 of Oikawa et al., a speed used to provide speech is determined by speed instruction generating unit 13 (see FIG. 2), and that any speed (normal, rapid speed 1, or rapid speed 2) may be used with any degree of importance, see Table 1 at cols. 3-4. In particular, Oikawa et al. does not teach, hint or suggest, in any manner whatsoever, obtaining user input regarding presentations rates for a portion of the input, or correlating content or properties of the portion with the user input presentation rates.

As evidence of a motivation to combine Richard et al. and Oikawa et al., the Examiner stated: "Therefore, it would have been obvious to one of ordinary skill at the time of invention to modify the system of Richard and implement associating playback rates based on specific categories as taught by Oikawa et al., for the purpose of ensuring that a user's preference for playback rates for a specific category of newspaper article is always maintained." Applicant

respectfully submits that Oikawa et al. does not teach using the same playback rate for a category of article, but teaches using a particular playback speed for a particular level of importance. Further, the user does not determine the level of importance, it is input with the data. Thus, the association of a level of importance and a playback rate would run at odds with Richards et al. that seeks to have the user assign a speech rate on an article by article basis, not on a content or properties basis. Hence, even if Richards et al. and Oikawa et al. were combined, one would not have the invention of claims 1 or 7.

Lastly, even if one were to use the Examiner's argument, that would still fall short of claims 1 or 7 because there would be no step of associating the affinity or aptitude with the presentation rates of the correlated portions.

In light of the above, Applicant respectfully submits that claims 1 and 7 are patentable over Richard et al. in view of Oikawa et al.

**Regarding claim 2.** Applicant respectfully submits that claim 2 depends from claim 1 and as such, is deemed patentable over Richard et al. in view of Oikawa et al. for the same reasons set forth above with respect to claim 1. In addition, Applicant respectfully submits there is nothing in Richard et al. or Oikawa et al. alone or in the combination of Richard et al. with Oikawa et al. that relates to a presentation rate which causes a portion of a media work to be skipped as required by claim 2. For example, Richard et al. teaches skipping a portion of a work based on text of a headline of an article of a newspaper and Oikawa et al. teaches skipping a portion of a work based on degree of importance which has nothing whatsoever to do with presentation rate.

In light of the above, Applicant respectfully submits that claim 2 is patentable over Richard et al. in view of Oikawa et al.

**Regarding claims 3 and 4.** Applicant respectfully submits that a combination of the teachings of Richard et al. and Oikawa et al. does not render claims 3 or 4 obvious because there is no teaching, motivation, or suggestion any where to combine them to provide the inventions of claims 3 or 4. In addition, even if one combined Richard et al. and Oikawa et al., one would not arrive at the inventions of claims 3 or 4.

In particular, claims 3 and 4 relate to a method of utilizing audience affinity or aptitude associated with content or properties to present a media work. In accordance with claims 3 and 4, the method includes detecting content or properties in a portion of a work, associating the audience affinity or aptitude associated with the detected content or properties with a presentation rate. There is nothing in Richard et al. or Oikawa et al. alone or in the combination of Richard et al. and Oikawa et al. that relates to associating audience affinity or aptitude associated with content or properties with presentation rates.

Richard et al. teaches that a speech rate is changed, if at all, independent of content. As one can readily appreciate from this, Richard et al. does not teach, hint or suggest associating a presentation rate of a portion of a newspaper with content or properties associated with audience affinity or aptitude. Instead, as set forth above, Richard et al. teaches that a speech rate is changed, if at all, independent of content.

Oikawa et al. does not teach, hint or suggest detecting content or properties in a work. In particular, Oikawa et al. teaches assigning an importance metric to content that is permanently assigned by a reviewer in an authoring step, and as such, Oikawa et al. does not teach, hint or suggest detecting content or properties in a work. As one can readily appreciate from this, Oikawa et al. does not teach, hint or suggest, in any manner whatsoever, associating a presentation rate of a portion of a newspaper with content or properties associated with audience affinity or aptitude.

Lastly, there is no teaching, motivation, or suggestion anywhere to combine Richard et al. or Oikawa et al. for the reasons set forth above with respect to claims 1 and 7. However, even if one did combine them, they still would not teach associating a presentation rate with detected content or properties because Richard et al. does not teach, hint or suggest doing this, and because Oikawa et al. teaches away from this.

In light of the above, Applicant respectfully submits that claims 3-4 are patentable over Richard et al. in view of Oikawa et al.

**Regarding claim 10.** Applicant respectfully submits that a combination of the teachings of Richard et al. and Oikawa et al. does not render claim 10 obvious because there is no teaching, motivation, or suggestion anywhere to combine them to provide the invention of

claim 10. In addition, even if one combined Richard et al. and Oikawa et al., one would not arrive at the invention of claim 10.

Applicant respectfully submits that Richard et al. and Oikawa et al. are both completely different from claim 10 which requires detecting information obtained from analyzing a media work other than content in a portion of a media work, and associating a presentation rate of the portion with the detected information other than text obtained from analyzing a media work. Such information is identified in the specification for example, and without limitation, at p. 96, line 20 - p. 97, line 13 such information could be speaker identification -- by voice or face; and at p. 97, line 26 - p. 98, line 17 such information could be a number of people in a camera view, number of objects in a scene, number of animals in a scene, and so forth). Thus, neither Richard et al. nor Oikawa et al. teach, hint or suggest, in any manner whatsoever, detecting information other than text obtained from analyzing a media work in a portion of a media work, and associating a presentation rate of the portion with the detected information. Lastly, there is no teaching, motivation, or suggestion anywhere to combine Richard et al. or Oikawa et al. However, even if one did combine them, they still would not teach detecting information other than text in a portion of a media work, and associating a presentation rate of the portion with the detected information because there is no teaching hint or suggestion to do this in either Richard et al. or Oikawa et al.

In light of the above, Applicant respectfully submits that claim 10 is patentable over Richard et al. in view of Oikawa et al.

**Regarding claim 11.** Applicant respectfully submits that a combination of the teachings of Richard et al. and Oikawa et al. does not render claim 11 obvious because there is no teaching, motivation, or suggestion anywhere to combine them to provide the invention of claim 11. In addition, even if one combined Richard et al. and Oikawa et al., one would not arrive at the invention of claim 11.

Applicant respectfully submits that Richard et al. and Oikawa et al. are both completely different from claim 11 which requires detecting information obtained from analyzing a media work other than content in a portion of a media work, and associating a presentation rate of the portion with the detected information other than text obtained from analyzing a media

work. In addition, in accordance with claim 11, such information comprise indicia of actions of objects. Such information is identified in the specification for example, and without limitation, at p. 96, line 20 - p. 97, line 13 such information could be speaker identification -- by voice or face; and at p. 97, line 26 - p. 98, line 17 such information could be a number of people in a camera view, number of objects in a scene, number of animals in a scene, and so forth). Thus, neither Richard et al. nor Oikawa et al. teach, hint or suggest, in any manner whatsoever, detecting information other than text obtained from analyzing a media work in a portion of a media work, and associating a presentation rate of the portion with the detected information. Lastly, there is no teaching, motivation, or suggestion anywhere to combine Richard et al. or Oikawa et al. However, even if one did combine them, they still would not teach detecting information other than text in a portion of a media work, and associating a presentation rate of the portion with the detected information because there is no teaching hint or suggestion to do this in either Richard et al. or Oikawa et al.

In light of the above, Applicant respectfully submits that claim 11 is patentable over Richard et al. in view of Oikawa et al.

**Issue 2:** **Whether claims 5-6 are patentable under 35 U.S.C. § 103(a) over Richard et al. (U.S. Patent No. 5,924,068) in view of Oikawa et al. (U.S. Patent No. 5,396,577) and well known prior art.**

**Reasons why claims 5 and 6 are separately patentable.**

Appellant respectfully submits that claim 5 is separately patentable because claim 5 relates to a method of presenting a media work by associating a presentation order with detected content or properties of the work, and reordering the portions according to the presentation order. Appellant respectfully submits that claim 6 is separately patentable because claim 6 relates to a method of presenting a media work by associating a presentation order and a presentation rate with detected content or properties of the work, and reordering the portions according to the presentation order.

**Richard et al.**

Richard et al. has been discussed above in responding to Issue 1.

**Oikawa et al.**

Richard et al. has been discussed above in responding to Issue 1.

**Combination of Richard et al. and Oikawa et al.**

The combination of Richard et al. and Oikawa et al. has been discussed above in responding to Issue 1. In addition Appellant respectfully submits that there is no reason, suggestion, or motivation in Richard et al. or Oikawa et al. or anywhere else that would have led one of ordinary skill in the art to combine Richard et al. and Oikawa et al to provide the methods of claims 5-6. In particular, the Examiner's evidence regarding a teaching, motivation, or suggestion for combining Richard et al. and Oikawa et al. to provide the method of claim 5 is: "Therefore, it would have been obvious to one of ordinary skill at the time of invention to modify the system of Richard et al. to implement reordering of the portions, for the purpose of allowing the user to hear the most desired portions first (i.e. weather before sports). (Emphasis added)" Applicant respectfully submits that this is a conclusory statement that does not fulfill the requirement (as set forth above) of the Federal Circuit. Further, because there is no teaching, motivation, or suggestion to combine Richard et al. and Oikawa et al. as asserted by the Examiner, Applicant respectfully submits that the Examiner's reasoning is based on improper hindsight. Still further, as will be set forth in detail below, even if one of ordinary skill in the art were to combine the teachings of Richard et al. and Oikawa et al., that one would not arrive at the method of claim 5. In particular, Richard et al. teaches a method for having a user hear certain sections first, and although Oikawa et al. enables a reordering by use of importance information, this would require the user to enter such information and for the apparatus of Richard et al. to be reworked.

In further particular, the Examiner's evidence regarding a teaching, motivation, or suggestion for combining Richard et al. and Oikawa et al. to provide the method of claim 6 is: "Therefore, it would have been obvious to one of ordinary skill at the time of invention to modify

the system of Richard and implement associating playback rates based on specific categories as taught by Oikawa et al., for the purpose of ensuring that a user's preference for playback rates for a specific category of newspaper is always maintained. (Emphasis added)" Applicant respectfully submits that this is a conclusory statement that does not fulfill the requirement (as set forth above) of the Federal Circuit. Further, because there is no teaching, motivation, or suggestion to combine Richard et al. and Oikawa et al. as asserted by the Examiner, Applicant respectfully submits that the Examiner's reasoning is based on improper hindsight. Still further, as will be set forth in detail below, even if one of ordinary skill in the art were to combine the teachings of Richard et al. and Oikawa et al., that one would not arrive at the method of claim 6. In particular, if one were to combine the teachings of Richard et al. and Oikawa et al., the resulting invention would require constantly having to input the importance information as taught by Oikawa et al.

**Regarding claim 5.** Applicant respectfully submits that a combination of the teachings of Richard et al. and Oikawa et al. does not render claim 5 obvious because there is no teaching, motivation, or suggestion anywhere to combine them to provide the invention of claim 5. In addition, even if one combined Richard et al. and Oikawa et al., one would not arrive at the invention of claim 5.

In particular, claim 5 relates to a method of presenting a media work that includes detecting content or properties in portions of the media work, associating a presentation order with the detected content or properties that is different from the order of detection, reordering the portions according to the presentation order, and presenting them in the presentation order. Richard et al. teaches receiving information from a newspaper transmitter regarding available sections, and storing the section identifiers in the order sent by the newspaper transmitter. The user indicates which of these sections to store by a setup procedure. See col. 7, lines 8-47. As set forth at col. 9, line 24 to col. 10, line 19, in the "basic mode" the sections are stored in the order specified during setup, i.e., the order in which the newspaper sent the available sections. As set forth at col. 15, lines 19-63, in an "advanced mode," during playback, the user can retrieve articles based on a keyword search. To do this, the user retrieves all articles containing

keywords. As set forth at col. 17, lines 37-53, a list of the articles is prepared, and the user can have them read in the order detected during the keyword search. As the Examiner can readily appreciate from this, Richard et al. does not teach, hint or suggest, in any manner whatsoever, associating a presentation order with the detected content or properties that is different from the order of detection. Further, Oikawa et al. does not teach, hint or suggest, in any manner whatsoever, associating a presentation order with the detected content or properties that is different from the order of detection. In addition, as set forth above, Oikawa et al. does not teach detecting content or properties in a work.

Even if one did combine Richards et al. and Oikawa et al., there would be no step of associating a presentation order with the detected content or properties that is different from the order of detection because neither Richard et al. nor Oikawa et al. teach, hint or suggest doing this.

As evidence of a motivation to combine Richard et al. and the prior art (the Examiner asserts: "Richard et al. does not specifically teach the reordering of the portions. However, reordering of presentation material was well known in the art.") the Examiner stated: "Therefore, it would have been obvious to one of ordinary skill at the time of invention to modify the system of Richard et al. to implement reordering of the portions, for the purpose of allowing the user to hear the most desired portions first (i.e. weather before sports)." Applicant respectfully submits that Examiner's assertion is based on improper hindsight. Further, any such reordering suggested by the Examiner would not necessarily be carried out in accordance with the requirements of claim 5. Lastly, Richard et al. teaches a method of searching so that the user could hear the newspaper articles in a desired order, so there would be no motivation to seek another method.

In light of the above, Applicant respectfully submits that claim 5 is patentable over Richard et al. in view of Oikawa et al.

**Regarding claim 6.** Applicant respectfully submits that a combination of the teachings of Richard et al. and Oikawa et al. does not render claim 6 obvious because there is no teaching, motivation, or suggestion anywhere to combine them to provide the invention of claim

6. In addition, even if one combined Richard et al. and Oikawa et al., one would not arrive at the invention of claim 6.

In particular, claim 6 relates to a method of presenting a media work that includes detecting content or properties in portions of the media work, associating a presentation order and a presentation rate with the detected content or properties, which presentation order is different from the order of detection, and presenting them in the presentation order at the presentation rate. Richard et al. teaches receiving information from a newspaper transmitter regarding available sections, and storing the section identifiers in the order sent by the newspaper transmitter. The user indicates which of these sections to store by a setup procedure. See col. 7, lines 8-47. As set forth at col. 9, line 24 to col. 10, line 19, in the "basic mode" the sections are stored in the order specified during setup, i.e., the order in which the newspaper sent the available sections. As set forth at col. 15, lines 19-63, in an "advanced mode," during playback, the user can retrieve articles based on a keyword search. To do this, the user retrieves all articles containing keywords. As set forth at col. 17, lines 37-53, a list of the articles is prepared, and the user can have them read in the order detected during the keyword search. As the Examiner can readily appreciate from this, Richard et al. does not teach, hint or suggest, in any manner whatsoever, associating a presentation order with the detected content or properties that is different from the order of detection or associating a presentation rate with the detected content or properties. Further, Oikawa et al. does not teach, hint or suggest, in any manner whatsoever, associating a presentation order with the detected content or properties that is different from the order of detection or associating a presentation rate with the detected content or properties. In addition, as set forth above, Oikawa et al. does not teach detecting content or properties in a work.

Even if one did combine Richards et al. and Oikawa et al., there would be no step of associating a presentation order with the detected content or properties that is different from the order of detection or associating a presentation rate with the detected content or properties because neither Richard et al. nor Oikawa et al. teach, hint or suggest doing this.

As evidence of a motivation to combine Richard et al. and the prior art (the Examiner asserts: "Richard et al. does not specifically teach the reordering of the portions.

However, reordering of presentation material was well known in the art.") the Examiner stated: "Therefore, it would have been obvious to one of ordinary skill at the time of invention to modify the system of Richard et al. to implement reordering of the portions, for the purpose of allowing the user to hear the most desired portions first (i.e. weather before sports)." Applicant respectfully submits that Examiner's assertion is based on improper hindsight. Further, any such reordering suggested by the Examiner would not necessarily be carried out in accordance with the requirements of claim 6. Lastly, Richard et al. teaches a method of searching so that the user could hear the newspaper articles in a desired order, so there would be no motivation to seek another method.

In light of the above, Applicant respectfully submits that claim 6 is patentable over Richard et al. in view of Oikawa et al.

**Issue 3:** **Whether claims 8-9 are patentable under 35 U.S.C. § 103(a) over Richard et al. (U.S. Patent No. 5,924,068) in view of Yumura et al. (U.S. Patent No. 5,752,228).**

**Reasons why claims 8-9 are separately patentable.**

Appellant respectfully submits that claim 8 is separately patentable because claim 8 relates to a method of presenting a media work that comprises accessing information identifying the media work and a time to retrieve the media work. Appellant respectfully submits that claim 9 is separately patentable because claim 9 relates to a method of presenting a media work that comprises accessing information identifying the media work and a time to retrieve the media work, and concatenating at least two altered media works.

**Richard et al.**

Richard et al. has been discussed above in conjunction with Issue 1.

**Yumura et al.**

As set forth at col. 1, lines 8-19, Yumura et al. teaches that: "The present invention relates to a speech synthesis apparatus for synthesizing speech on the basis of the text data at a speed that can finish reading out a text within a fixed time ... Further, the present invention relates to a read out time calculating apparatus for calculating time necessary for a reader to finish reading out a text, according to a speaking speed extracted from the reader's speech data ... ." Next, at col. 1, lines 21-26, Yumura et al. teaches that: "The time permitted to read out a manuscript or to narrate is limited within an announcement time prepared for each speaker in a lecture, speech or the like, within the time a title is displayed on a screen, within a prelude or interlude being played, or within the time a picture relating to the contents of a story is displayed on the screen." Next, at col. 1, lines 57-63, Yumura et al. teaches that: "It is an object of the invention to provide a speech synthesis apparatus and a medium on which is recorded a computer program for reading out a text in place of a reader where the synthesized speech at such a speed that can finish reading out the text takes the place of the reader ... ." Next, at col. 1, line 58 to col. 2, line 3, Yumura et al. teaches that: "Another object of the invention is to provide a speech synthesis apparatus ... for reading out a text in place of a reader where the synthesized speech is extremely like speech by the reader thereby lightening the burden of completing the text by the reader within a fixed time." Next, at col. 2, lines 9-14, Yumura et al. teaches that: "Yet another object of the invention is to provide a read out time calculating apparatus and a medium on which is recorded a computer program for calculating the time to finish reading out the text without actually reading out the text, thereby lightening the burden of the reader in completing the text." Next, at col. 2, lines 13-31, Yumura et al. teaches that:

> A speech synthesis apparatus ... of the invention ... for reading out a text in place of a reader, calculates time to read out the text at a prescribed read out speed when the fixed time to read out the text is set, determines the read out speed which makes the calculated time agree with the set time on the basis of the text data, then synthesizes speech at the determined read out speed. A user judges whether the read out speed which enables the text to be read out within the set time is appropriate for sufficiently transmitting the contents of the text on listening to the synthesized speech.

When judging the contents being sufficiently transmitted, the user makes the prescribed reader read the text at the determined speed without changing the contents, but when judging the contents being insufficiently transmitted, the user adjusts the contents of the text. Consequently, it is unnecessary to actually read out the text for judging whether the reading speed is appropriate.

Next, at col. 2, lines 55-62, Yumura et al. teaches that:

A read out time calculating apparatus ... of the invention, on which is recorded a computer program for calculating time to finish reading out a text, calculates the time to finish reading out text data at a read out speed being set, then outputs the calculated time. A user deletes or supplements the contents of the text on referring to the output time so that the time necessary for finishing the read out the text is nearly at a prescribed time.

Next, at col. 3, lines 43-67, Yumura et al. teaches that:

FIG. 1 is a block diagram of a speech synthesis apparatus of the invention. ... A morphological analysis unit 2 cuts out the text data sentence by sentence ... referring to a morpheme dictionary 3, then morphologically analyzes the sentence to attach a part of speech and accent data thereto. ... The morphological analysis unit 2 extracts punctuation of a clause and accent phrases, and attaches pause data necessary to put a pause in reading. The morphological analysis unit 2 further performs a phonemic language processing on the text data to add focus data to a part necessary to be phonetically emphasized and to attach speed control data according to presence or absence of the focus data. A reference read out time calculating unit 4 changes the length of a mora (tactus) which is a time unit, corresponding to speaking time of a normal syllable, on a time scale of a speech waveform in order to read out the focused part of the text data slowly. Then, the reference read out time calculating unit 4 adds up the read out time of each sentence of the text at a reference read out speed having a reference read out speed parameter to calculate the reference read out time of the whole text.

Next, at col. 4, lines 1-48, Yumura et al. teaches that:

A read out time setting device 5 is composed of a ten-key pad or the like for setting a time to finish reading out a text. A read out speed determining unit 6 determines a read out speed parameter which makes the reference read out time agree with the set read out time on comparing the read out time set by the read out time setting device 5 with the reference read out time calculated by the reference read out time calculating unit 4.

A speech database 7 stores unit waveform signals of the text data as data for the speech synthesis obtained by dividing the text data into units which are suitable for speaking for the speech synthesis ... on the basis of the phonological analysis or the like, thereby enabling the text to be read out in a way as natural as possible. The speech database 7 further stores speech characteristic data of a reader preliminarily extracted from a frequency spectrum of speech data of the reader obtained by speaking a prescribed word, sentence or the like.

A speech synthesis unit 8 reads out the data for performing the speech synthesis for the text data, and the speech characteristic in order to perform a waveform signal processing for linking the data for the speech synthesis of every unit having the reader's speech characteristic, thereby enabling the text data to be smoothly read out. Then the speech synthesis unit 8 outputs the synthesized speech from a speaker 9 as if the reader is reading out the text.

... In the figure, numeral 11 designates a speech input device 11 such as a microphone. A read out speed extracting unit 12, which stores speech data of a prescribed word, sentence or the like spoken at a reference speed, extracts a parameter of the read out speed of a reader relative to the reference read out speed on comparing the speech data of the prescribed word or sentence spoken by the reader and input through the speech input device 11 with the speech data at the reference read out speed.

A read out time adjusting unit 13 adjusts the reference read out time calculated by the reference read out time calculating unit 4 on the basis of the read out speed parameter extracted by the read out speed extracting unit 12 to calculate the read out time of the text by the reader. The read out time adjusting unit 13 displays the read out time of the reader on a monitor 14.

Next, at col. 4, lines 55-67, Yumura et al. teaches that:

This modified embodiment differs from the abovementioned embodiment in setting the read out speed but not

extracting that from the speech data input through the microphone. Therefore, the apparatus is provided with a read out speed setting device 15 and a read out time calculating unit 16. The read out time calculating unit 16 changes the length of a mora (tactus) which is a time unit, corresponding to speaking time of a normal syllable, on a time scale of a speech waveform in order to read out the focused part of the text data slowly. Then, the read out time calculating unit 16 adds up the read out time of each sentence of the text at the set read out speed up to calculate the read out time of the whole text.

Next, at col. 5, lines 1-67, Yumura et al. teaches that:

The procedure of reading out a Japanese text by the speech synthesis apparatus of the invention instead of a reader will be explained according to flowcharts in FIG. 4 and FIG. 5.

When text data is input through the text input device 1 (S1), the morphological analysis unit 2 cuts out one sentence from the input text data (S2). Then, the morphological analysis unit 2 analyzes the text data into morphemes to attach a part of speech and accent data to each morpheme referring to the morpheme dictionary 3 (S3). The morphological analysis unit 2 further attaches the "yomi" to each morpheme. The morphological analysis unit 2 extracts a clause, an accented phrase to attach pause data to a part necessary to put a pause in reading (S5).

...

The morphological analysis unit 2 performs a phonemic language processing on the text data to add focus data to a part necessary to be phonetically emphasized and to attach speed control data to read the focus data added part slowly according to a part with the focus data being attached (S6). The reference read out time calculating unit 4 changes the length of a mora so as to read out the focused part of the text data slowly (S7). Then, the reference read out time calculating unit 4 calculates reference read out time of one sentence at a reference read out speed (S8), and adds up the read out time of each sentence at the reference read out speed to calculate the reference read out time of the whole text (S9).

On the other hand, when time to finish reading out the text is set through the read out time setting device 5, the read out speed determining unit 6 determines a read out speed parameter which makes the reference read out time agree with the set read out time.

In other words, the read out speed parameter which enables the text to be read within the set time is set, on comparing the read out time set by the read out time setting device 5 with the reference read out time calculated by the reference read out time calculating unit 4 (S11).

By performing the above-mentioned steps S2-S7, the "yomi", pause data, and speed control data depending on presence and absence of the focus data are attached to each mora of each sentence (S12-S16), and the length of each mora is changed (S17) on the basis of the calculated read out speed parameter as above. The speech synthesis unit 8 synthesizes speech at the read out speed which enables the text data to be read within the set time on the basis of the adjusted parameters according to the read out time parameter which enables the text data to be read within the set time and on the basis of the stored speech characteristic data of the reader in the speech database 7 (S18), and outputs the synthesized speech from the speaker 9 (S19). By repeating the above processing for every sentence, the synthesized speech of the whole text is output.

Next, at col. 6, lines 1-9, Yumura et al. teaches that:

A user judges whether the read out speed which enables the text to be read out within the set time is appropriate for sufficiently transmitting the contents of the text on listening to the synthesized speech, reading out the text. When judging the contents being sufficiently transmitted, the reader reads the text at the determined speed without changing the contents, but when judging the contents being insufficiently transmitted, the user deletes or summarizes the contents of the text.

Next, at col. 6, lines 25-65, Yumura et al. teaches that:

The procedure of calculating the read out time by the read out time calculating apparatus of the invention will be explained according to a flowchart in FIG. 9.

When text data is input through the text input device 1 (S21), the morphological analysis unit 2 cuts out one sentence from the input text data (S22). Then, the morphological analysis unit 2 analyzes the text data into morphemes to attach a part of speech and accent data to each morpheme referring to the morpheme

dictionary 3 (S23). The morphological analysis unit 2 further attaches the "yomi" to each morpheme (S24). The morphological analysis unit 2 extracts a clause, an accented phrase to attach pause data to a part necessary to put a pause in reading (S25).

The morphological analysis unit 2 performs a phonemic language processing on the text data to add focus data to a part necessary to be phonetically emphasized and to attach speed control data to read the focus data attached part slowly according to a part with the focus data being attached (S26). The reference read out time calculating unit 4 changes the length of a mora so as to read out the focused part of the text data slowly (S27). Then, the reference read out time calculating unit 4 calculates reference read out time of one sentence at a reference read out speed (S28), and adds up the read out time of each sentence at the reference read out speed to calculate the reference read out time of the whole text (S29).

On the other hand, when the reader's speech is input through the speech input device 11 (S30), the read out speed extracting unit 12 extracts a parameter of the read out speed of the reader relative to the reference read out speed on comparing the speech data of the prescribed word or sentence spoken by the reader and input through the speech input device 11 with the speech data at the reference read out speed (S31). The read out time adjusting unit 13 adjusts the reference read out time calculated by the reference read out time calculating unit 4 on the basis of the read out speed parameter extracted by the read out speed extracting unit 12 (S32) to calculate the read out time of the text by the reader. The read out time adjusting unit 13 displays the read out time of the reader on a monitor 14 (S33).

**Combination of Richard et al. and Yumura et al.**

Appellant respectfully submits that there is no reason, suggestion, or motivation in Richard et al. or Yumura et al. or anywhere else that would have led one of ordinary skill in the art to combine Richard et al. and Yumura et al. to provide the methods of claims 8-9. In particular, as discussed above, Richard et al. relates to an apparatus that receives an electronic version of a newspaper, stores pre-selected types of articles, "speaks" the headlines of stored articles, and "speaks the articles in response to user input. This is contrasted with Yumura et al. which relates to a speech synthesis apparatus that "speaks" input text in a time limit. To utilize

the apparatus disclosed in Yumura et al., a calculation of the length of time of spoken text is determined. If the length of time is too long, the speech is speeded up or some of the text is deleted. There is no reason to combine these references because they are so different. In particular, Richard et al. teaches an elaborate methodology for a user to determine which selections are to be skipped by listening to headlines for various sections of the electronic newspaper, whereas Yumura et al. provides an output in a time provided by user input. There is no teaching or suggestion that the person listening to the newspaper needs to have the articles read in a certain time frame. However, even if they did, one would not arrive at the methods of claims 8-9 which require creating an altered work. In addition, the Examiner's evidence regarding a teaching, motivation, or suggestion for combining Richard et al. and Yumura et al. to provide the methods of claims 8-9 is: "Therefore, it would have been obvious to one of ordinary skill at the time of invention to modify the system of Richard et al. to specifically provide for a reference read out rate as taught by Yumura, for the purpose of providing a reference speed reflective of normal speech of which the varied rate is based on, ensuring that the synthetic speech is intelligible and natural. (Emphasis added)" Applicant respectfully submits that this is a conclusory statement that does not fulfill the requirement (as set forth above) of the Federal Circuit. Further, because there is no teaching, motivation, or suggestion to combine Richard et al. and Yumura et al. as asserted by the Examiner, Applicant respectfully submits that the Examiner's reasoning is based on improper hindsight. Still further, as will be set forth in detail below, even if one of ordinary skill in the art were to combine the teachings of Richard et al. and Yumura et al., that one would not arrive at the methods of claims 8-9.

**Regarding claim 8.** Applicant respectfully submits that a combination of the teachings of Richard et al. and Yumura et al. al. does not render claim 8 obvious because there is no teaching, motivation, or suggestion anywhere to combine them to provide the invention of claim 8. In addition, even if one combined Richard et al. and Yumura et al., one would not arrive at the invention of claim 8.

In particular, claim 8 relates to a method of presenting a media work that includes accessing information identifying the media work and a time to retrieve the media work, retrieving the identified media at the time, accessing presentation rate information to obtain a

new presentation rate, and altering the media work to create an altered work having the new presentation rate.

Applicant respectfully submits that Richard et al. does not teach, hint or suggest, in any manner whatsoever, a step of accessing information identifying a media work as required by claim 8. This can be appreciated from the following, see Richard et al. at col. 8, lines 4-27:

> Referring to FIG. 1, once the section list and the associated section bits have been stored in memory 141, the newsreader 100 is ready to receive the electronic newspaper. In the exemplary embodiment, the receiver 130 receives the electronic edition of the newspaper in the early morning hours before dawn. The receiver 130 contacts the transmitter 120 and initiates the transfer of the electronic edition of the newspaper to the newsreader 100. FIG. 9 shows hardware elements within the receiver 120 which are used for the automatic downloading of the electronic edition of the newspaper. As shown in FIG. 9, the receiver 130 includes a timer 1310. The timer 1310 determines whether it is time to call the electronic news preparer. At the appropriate time (some time in the early morning) the timer 1310 instructs the modem 1320 to call the transmitter. As described above, other communications devices may be used. For example, if the electronic edition of the newspaper is received through AM/FM SCA broadcast, the timer would enable a tuner/demodulator circuit (instead of modem 1320) to begin receiving the electronic edition of the newspaper. In addition, the electronic edition of the newspaper may be provided from another source such as a cable company which is provided the electronic edition by the electronic news provider. (Emphasis added)

As the Examiner can see from this, Richard et al. does not teach or disclose a step of "accessing information identifying a media work" as required by claim 8.

In addition, Applicant respectfully submits that, although Richard et al. teaches (see the quote set forth above) using a timer that causes a receiver to retrieve an electronic newspaper, this is not a step of accessing a time to retrieve the media work as required by claim 8.

In further addition, Applicant respectfully submits that Richard et al. does not teach or disclose altering a presentation rate of a media work to create an altered work as

required by claim 8. This can be appreciated from the following quote, see Richard et al. at col. 19, lines 9-12:

> Speech Rate
> The text-to-speech converter 170 has an average default speech rate is 170 words per minute. The user can select an average speech rate from 120 to 240 words per minute.

As the Examiner can see from this, Richard et al. teaches using an average speech rate to synthesize speech from text. Thus, Richard et al. does not teach or disclose a step of altering the presentation rate of the media work to create an altered work as required by claim 8.

Applicant respectfully submits that Yumura et al. teaches a speech synthesis apparatus for synthesizing speech on the basis of the text data at a speed that can finish reading out a text within a fixed time. In addition, Yumura et al. teaches a read out time calculating apparatus for calculating the time to finish reading out the text at a prescribed read out speed when the fixed time to read out the text is set. As such, Yumura et al. does not teach, hint or suggest, in any manner whatsoever, accessing information identifying the media work and a time to retrieve the media work as required by claim 8.

Even if one did combine Richards et al. and Yumura et al., there would be no step of accessing information identifying the media work and a time to retrieve the media work because neither Richard et al. nor Yumura et al. teach, hint or suggest doing this.

As evidence of a motivation to combine Richard et al. and Yumura et al. the Examiner stated: "Therefore, it would have been obvious to one of ordinary skill at the time of invention to modify the system of Richard et al. to specifically provide for reference read out rate as taught by Yumura et al., for the purpose of providing a reference speed reflective of normal speech of which the varied rate is based on, ensuring that the synthetic speech is intelligible. Applicant respectfully submits that Examiner's assertion would not supply the step of accessing information identifying the media work and a time to retrieve the media work in accordance with the requirements of claim 8. Lastly, Examiner's assertion would not be true since, in accordance with Yumura et al. at col. 6, lines 1-9, a human must listen to the speech to determine whether it is being "sufficiently transmitted."

In light of the above, Applicant respectfully submits that claim 8 is patentable over Richard et al. in view of Yumura et al.

**Regarding claim 9.** Applicant respectfully submits that claim 9 depends from claim 8 and as such, is deemed patentable over Richard et al. in view of Yumura et al. for the same reasons set forth above with respect to claim 8. In addition, Applicant respectfully submits that Richard et al. does not teach or disclose a step of concatenating several altered media works to form a concatenated media work as required by claim 9. This can be appreciated from the following quote, see Richard et al. at col. 9, line 63 to col. 10, line 2:

> To playback the stored articles, the user operates the playback controls 162, shown in detail in FIG. 13. The user first presses the Sections button 511. The headlines of the articles in the current section are consecutively read to the user. In addition, the current section is displayed on LCD 430 (shown in FIG. 4). Once the headlines for the current section are read, the newsreader automatically begins reading the headlines for the next section.

As the Examiner can see from this, Richard et al. does not teach or suggest creating an altered work, let alone several altered works. Thus, Richard et al. does not teach a step of concatenating several altered media works to form a concatenated media work as required by claim 9.

In light of the above, Applicant respectfully submits that claim 9 is patentable over Richard et al. in view of Yumura et al.

**Issue 4:** **Whether claims 12-13 are patentable under 35 U.S.C. § 103(a) over Yumura et al. (U.S. Patent No. 5,752,228) in view of Richard et al. (U.S. Patent No. 5,924,068).**

**Reasons why claims 12-13 are separately patentable.**

Appellant respectfully submits that claim 12 is separately patentable because claim 12 relates to a method of determining the duration of an altered media work. In addition, Applicant respectfully submits that claim 13 is separately patentable because claim 12 relates to a

method of determining the duration of an altered media work that comprises excising segments from the media work having a presentation rate that exceeds a predetermined threshold.

**Richard et al.**

Richard et al. has been discussed above in responding to Issue 1.

**Yumura et al.**

Yumura et al. has been discussed above in responding to Issue 3.

**Combination of Yumura et al. and Richard et al.**

The combination of Richard et al. and Yumura et al. has been discussed above in responding to Issue 3. In addition, the Examiner's evidence regarding a teaching, motivation, or suggestion for combining Richard et al. and Yumura et al. to provide the methods of claims 12-13 is: "Therefore, it would have been obvious to one of ordinary skill at the time of invention to modify the system of Yumura to implement altering the presentation rate of the read out material as taught by Richard for the purpose of providing varied presentation of the read out text based on a user's preference to have the information presented slowly if important or quickly if non-essential. (Emphasis added)" Applicant respectfully submits that this is a conclusory statement that does not fulfill the requirement (as set forth above) of the Federal Circuit. Further, because there is no teaching, motivation, or suggestion to combine Richard et al. and Yumura et al. as asserted by the Examiner, Applicant respectfully submits that the Examiner's reasoning is based on improper hindsight. Still further, as will be set forth in detail below, even if one of ordinary skill in the art were to combine the teachings of Richard et al. and Yumura et al., that one would not arrive at the methods of claims 12-13.

**Regarding claim 12.** Applicant respectfully submits that a combination of the teachings of Yumura et al. and Richard et al. does not render claim 12 obvious because there is no teaching, motivation, or suggestion anywhere to combine them to provide the invention of

claim 12. In addition, even if one combined Yumura et al. and Richard et al., one would not arrive at the invention of claim 12.

In particular, claim 12 relates to a method of determining the duration of an altered media work, and in accordance with claim 12, the media work is segmented into segments having a single presentation rate, lengths of the segments are determined, the duration of the segments are determined, and summed.

Applicant respectfully submits that Yumura et al. does not teach, hint or suggest, in any manner whatsoever, a step of segmenting the media work into segments having a single presentation rate. In fact, as set forth at col. 5, lines 1-67 of Yumura et al., "Then, the reference read out time calculating unit 4 calculates reference read out time of one sentence at a reference read out speed (S8), and adds up the read out time of each sentence at the reference read out speed to calculate the reference read out time of the whole text (S9)." Thus, Yumura et al. teaches cutting the text data into sentences, performs morphological analysis, and determines the time to read the sentence at a fixed speed, i.e., a speed "so as to read out the focused part of the text data slowly."

Applicant respectfully submits that Richard et al. does not teach performing a duration calculation for use in computing a duration of presentation of the work. This can be appreciated from the following quote, see Richard et al. at col. 14, lines 36-43:

> The duration calculation module 723 computes the length of each phoneme segment based on several observations about the segment and its environment. For instance, vowel phonemes are generally longer than consonant phonemes. Also, phonemes which precede the vowel in a stressed word are longer than the same phonemes in a non-stressed word. Several similar rules are applied to each segment to calculate the final duration of each segment.

As the Examiner can appreciate from this quote, Richard et al. teaches performing a duration calculation merely to create spoken output. In other words, Richard et al. teaches performing a duration calculation to enable speech output to sound as much like a natural person as possible, and not like a machine. For example, Richard et al. teaches that a phoneme which

precedes a vowel in a stressed word has longer duration than the same phoneme in a non-stressed word because, otherwise, the speech would sound "machine-like."

Thus, although Richard et al. teaches determining durations of phoneme segments, this does not in any way teach, hint or suggest segmenting a media work into segments having a single presentation rate, determining the length of the segments, computing the duration of the segments after applying a presentation rate.

Lastly, there is no teaching, motivation, or suggestion anywhere to combine Yumura et al. and Richard et al. However, even if one did combine them, they still would not teach segmenting a media work into segments having a single presentation rate, determining the length of the segments, computing the duration of the segments after applying a presentation rate because there is no teaching hint or suggestion to do this in either Yumura et al. or Richard et al.

As evidence of a motivation to combine Yumura et al. and Richard et al. the Examiner stated: "Therefore, it would have been obvious to one of ordinary skill at the time of invention to modify the system of Yumura et al. to implement altering the presentation rate of the read out material as taught by Richard et al., for the purpose of providing varied presentation of the read out text based on a user's preference to have the information presented slowly if important or quickly if non-essential. Applicant respectfully submits that Examiner's assertion would not supply the step of segmenting the media work into segments having a single presentation rate in accordance with the requirements of claim 12.

In light of the above, Applicant respectfully submits that claim 12 is patentable over Yumura et al. in view of Richard et al.

**Regarding claim 13.** Applicant respectfully submits that claim 13 depends from claim 12 and as such, is deemed patentable over Yumura et al. in view of Richard et al. for the same reasons set forth above with respect to claim 12. In addition, Applicant respectfully submits that neither Yumura et al. nor Richard et al. teach, hint or suggest, in any manner whatsoever, a step of excising segments from the media work having a presentation rate that exceeds a predetermined threshold as required by claim 13. This can be appreciated from Yumura et al. at col. 2, lines 13-31, which teaches: "A speech synthesis apparatus ... of the

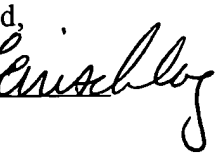
invention ... for reading out a text in place of a reader, calculates time to read out the text at a prescribed read out speed when the fixed time to read out the text is set, determines the read out speed which makes the calculated time agree with the set time on the basis of the text data, then synthesizes speech at the determined read out speed. A user judges whether the read out speed which enables the text to be read out within the set time is appropriate for sufficiently transmitting the contents of the text on listening to the synthesized speech. When judging the contents being sufficiently transmitted, the user makes the prescribed reader read the text at the determined speed without changing the contents, but when judging the contents being insufficiently transmitted, the user adjusts the contents of the text." Thus, since Yumura et al. is concerned with speeding up the reading out of text, Yumura et al. does not teach excising segments from a media work having a presentation rate that exceeds a predetermined threshold. In further addition, Yumura et al. teaches slowing the speaking rate if the text cannot be properly understood. In further addition , Yumura et al. teaches deleting sections if they cause the total read-out time of text to exceed a predetermined time, and to excise segments having a presentation rate that exceeds a predetermined threshold.

In light of the above, Applicant respectfully submits that claim 13 is patentable over Yumura et al. in view of Richard et al.

In light of the above, Appellant respectfully submits that claims 1-13 are patentable. Accordingly, Appellant respectfully requests that the Examiner's rejections be reversed, and the application be allowed at the earliest opportunity.

Respectfully submitted,

Date: December 24, 2003

By 

Michael B. Einschlag
Reg. No. 29,301
(650) 949-2267
25680 Fernhill Drive
Los Altos Hills, California 94024

MBE/pc

**Appendix.**

Copy of Claims 1-13 Involved in the Appeal

1. A method for inferring audience affinity or aptitude with regard to content or properties of portions of a media work which comprises:

presenting the media work to an audience;

obtaining user input regarding presentation rates for the portions of the media work;

correlating the content or properties of the portions with the presentation rates; and;

associating audience affinity or aptitude with the presentation rates for the correlated content or properties.

2. The method of claim 1 wherein the presentation rates include a rate which causes a portion to be skipped.

3. A method of utilizing audience affinity or aptitude associated with content or properties to present a media work which comprises:

detecting the content or properties in a portion of the media work;

associating the audience affinity or aptitude associated with the detected content or properties with a presentation rate for the portion; and

presenting the portion at the presentation rate.

4. The method of claim 3 wherein associating includes accepting user input to determine the presentation rate.

5. A method of presenting a media work which comprises:

detecting content or properties in portions of the media work;

associating a presentation order with the detected content or properties that is different from the order of detection;

reordering the portions according to the presentation order; and

presenting the media work in accordance with the presentation order.

6. A method of presenting a media work which comprises:

detecting content or properties in portions of the media work;

associating a presentation order with the detected content or properties that is different from the order of detection; and

presenting the media work in accordance with the presentation order;

wherein the step of associating further comprises associating a presentation rate of the portion with the detected content or properties; and the step of presenting comprises presenting the media work in accordance with the presentation order and the presentation rates.

7. A method of testing aptitude of an audience for content or properties of portions of a media work which comprises:

presenting the media to the audience;

obtaining user input regarding presentation rates for the portions of the media work; and

correlating the presentation rates with the aptitude for the content or properties of the portions.

8. A method of presenting a media work having a presentation rate which comprises:

accessing information identifying the media work and a time to retrieve the media work;

retrieving the identified media work at the time;

accessing presentation rate information to obtain a new presentation rate for use in altering the media work; and

altering the media work to create an altered work having the new presentation rate.

9. The method of claim 8 which further comprises:

concatenating at least two altered media works to form a concatenated media work; and

presenting the concatenated media work.

10. A method of presenting a media work which comprises:

detecting media work content properties in a portion of the media work;

associating a presentation rate of the portion with the detected media work content properties; and

presenting the portion at the presentation rate;

wherein the presentation rates provide a substantially uniform rate of content presentation.

11. A method of presenting a media work which comprises:

detecting media work content properties in a portion of the media work;

associating a presentation rate of the portion with the detected media work content properties;

presenting the portion at the presentation rate; and

wherein the media work content properties comprise indicia of actions of objects.

12. A method of determining the duration of an altered media work having a presentation rate of one or more of its segments that differs from that of a media work used to create the altered media work, which method comprises:

segmenting the media work into segments having a single presentation rate;

determining the length of the segments of the media work;

computing the duration of the segments of the media work after application of the presentation rate; and

summing the durations to determine the duration of the altered media work.

13. The method of claim 12 which further comprises:

excising segments from the media work having a presentation rate that exceeds a predetermined threshold.